Bayesian Hierarchical Models to Augment the Mediterranean Forecast System Consolidating Results and Quantifying Impacts

L. Mark Berliner
Department of Statistics
Ohio State University
1958 Neil Avenue
Columbus, OH, USA 43210-1247
phone:614-292-0291 fax:614-292-2096 email: mb@stat.ohio-state.edu

Emanuele Di Lorenzo
School of Earth and Atmospheric Sciences
Georgia Institute of Technology
311 Ferst Drive
Atlanta, GA, USA 30332-0340
phone: 404-894-3994 fax: 404-894-5638 email: edl@eas.gatech.edu

Christopher K. Wikle
Department of Statistics
University of Missouri-Columbia
146 Middlebush
Columbia, MO, USA 65211
phone:573-882-9659 fax:573-884-5524 email: wikle@stat.missouri.edu

Ralph F. Milliff
Colorado Research Associates Division
NorthWest Research Associates Inc.
3380 Mitchell Lane
Boulder, CO, USA 80301-5410
phone:303-415-9701 x216 fax:303-415-9702 email: milliff@cora.nwra.com

Award Number: N00014-07-1-0512
http://www.cora.nwra.com/MedBhm (password protected)

LONG-TERM GOALS

The long term goal of this interdisciplinary research program continues to be the demonstration of Bayesian Hierarchical Model (BHM) utility in several aspects of operational ocean forecasting.

The specific goals in the current phase of the research are: 1) publications of results from ensemble ocean forecast experiments driven by the surface vector wind (SVW) BHM; b) consolidating impacts results in MFS reforecast experiments for the time-dependent error-covariance BHM; and c) running the first multi-model and multi-parameter super-ensemble BHM experiments.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **30 SEP 2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Bayesian Hierarchical Models to Augment the Mediterranean Forecast System Consolidating Results and Quantifying Impacts**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Ohio State University,Department of Statistics,1958 Neil Avenue,Columbus,OH,43210-1247**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

OBJECTIVES

Research objectives leading to the publication of manuscripts regarding SVW-BHM include:

1) constructing 3 appendices for Milliff et al (2009) to: a) demonstrate a systematic approach to (future) process model development; b) document, in probability model notation, the complete SVW-BHM, as well as expressions for the full conditional distributions; and c) document the SVW-BHM hyperprior specifications;
2) Re-writing the text and updating figures for Bonazzi et al. (2009); and
3) incorporating co-author final edits for Milliff et al. (2009) and Bonazzi et al. (2009).

Research objectives leading to consolidation of results for the time-dependent error-covariance BHM include:

1) supplying an anomaly-only data stage version of the error-covariance BHM to MFS for reforecast experiments; and
2) interpreting reforecast results, and iterating with MFS for future reforecast experimental design.

Research objectives leading to the first multi-model and multi-parameter super-enesmble BHM runs include:

1) evaluating preliminary experiments based on a Levantine Intermediate Water (LIW) formation rate target process;
2) re-organizing the target process to focus on temperature (*T*) and salinity (*S*) profile evolution at 2 locations in the region of LIW formation (Rhodes Gyre); and
3) providing *T(z,t)* and *S(z,t)* data files (.mat) to Mark Berliner for preliminary modelling.

APPROACH

Appendices and updates for the BHM-SVW papers:

The sample BHM model equation and the expression for the full conditional distribution for the zonal wind component $\boldsymbol{u}_t$ are provided here, where $\boldsymbol{u}$ is an *n-vector,* spanning all *n* locations on the target grid. The process model form (derived from the Rayleigh Friction Equations) is:

$$\mathbf{u}_t|\mathbf{P}_t, a_{1,1}, a_{1,2}, \beta_t^u, \sigma_u^2 \sim N(a_{1,1}D_y\mathbf{P}_t + a_{1,2}D_x\mathbf{P}_t + \mathcal{W}\beta_t^u, \sigma_u^2\mathbf{I})$$

where there are terms in the Normal distribution specification (right hand side) for pressure gradient effects and a scale-selective nested wavelet summation that proved useful in a process model for tropical winds in Wikle et al. (2001). In keeping with a key advantage of BHM, the uncertainty in the zonal wind distribution is modeled in the variance-covariance term. If we let $H_u = a_{1,1}D_y + a_{1,2}D_x$ then the full conditional distribution is:

$$
\begin{aligned}
[\mathbf{u}_t|\cdot] \;\; &\propto \;\; [\mathbf{D}_t^s|\mathbf{u}_t,\sigma_s^2]\,[\mathbf{D}_t^{eu}|\mathbf{u}_t,\sigma_e^2]\,[\mathbf{u}_t|\mathbf{P}_t,H_u,\beta_t^u,\sigma_u^2\,] \\
&\propto \;\; exp\left\{-\frac{1}{2}\left[\frac{(\mathbf{D}_t^{su}-K_t^s\mathbf{u}_t)'(\mathbf{D}_t^{su}-K_t^s\mathbf{u}_t)}{\sigma_s^2}\right.\right. \\
&\qquad + \frac{(\mathbf{D}^{eu}-K_t^e\mathbf{u}_t)'(\mathbf{D}^{eu}-K_t^e\mathbf{u}_t)}{\sigma_e^2} \\
&\qquad \left.\left. + \;\; \frac{(\mathbf{u}_t-H_u\mathbf{P}_t-\mathcal{W}\beta_t^u)'(\mathbf{u}_t-H_u\mathbf{P}_t-\mathcal{W}\beta_t^u)}{\sigma_u^2}\right]\right\}.
\end{aligned}
$$

As noted above, this can be shown to be a Gaussian Distribution $\boldsymbol{u}_t|\cdot \sim N(\boldsymbol{Ab},\boldsymbol{A})$, where

$$
\begin{aligned}
\mathbf{A} \;\; &= \;\; \left[\frac{K_t^{s\prime}K_t^s}{\sigma_s^2}+\frac{K^{e\prime}K^e}{\sigma_e^2}+\frac{\mathbf{I}}{\sigma_u^2}\right]^{-1} \\
\boldsymbol{b} \;\; &= \;\; \left[\frac{\mathbf{D}_t^{su\prime}K_t^s}{\sigma_s^2}+\frac{\mathbf{D}^{eu\prime}K^e}{\sigma_e^2}+\frac{(H_u\mathbf{P}_t+\mathcal{W}\beta_t^u)'}{\sigma_u^2}\right]'.
\end{aligned}
$$

Expressions of these forms for data stage and process model terms, and the implied full conditional distributions of BHM-SVW are included in Appendix 2 of Milliff et al. (2009). The hyperpriors for terms that are specified (i.e. rather than endowed with distributions) in the full conditional distributions are included in Appendix 3.

The companion paper, Bonazzi et al. (2009) was updated and re-written by Prof. Nadia Pinardi during her annual summer visit to Boulder, CO (CoRA offices) this August. Both papers are undergoing final edits by co-authors before submission for publication this month.

Iterating Time-Dependent Error Covariance BHM based on impacts:

A new round of reforecast experiments for the Gulf of Lyon region (MFS region 3) have been run to test time-dependent error covariance matrices that are the posterior means of the Error-Covariance BHM. Earler experiments based on a version of the time-dependent error covariance BHM using the quality controlled misfit ***d*** and anomaly ***q*** data provided by MFS did not lead to significant improvements in comparisons with Argo data. The comparisons were noisy, suggesting that very localized and short timescale variability was influencing the time-dependent error covariance in ways that were not useful to the forecast system.

The new version of the error covariance BHM uses only ***q*** information in the data stage. The ***q*** inputs are more smoothly varying in time than ***d***, and they exhibit a clear seasonal cycle. Time series of error covariance matrices were provided to MFS for "***q-only***" reforecast experiments.

Identifying T(z,t), S(z,t) Target Processes for the Super-Ensemble BHM

The LIW formation rate time series from analysis (SYS3a2) and simulations from two MFS ocean models (OPA and NEMO) did not lead to informative posterior distributions in preliminary experiments with the super-ensemble BHM, as designed and run by Prof. Mark Berliner (i.e. as in

Berliner and Kim, 2008). There are ambiguities in the definition of LIW, and the model processes leading to formation, that are larger than the signal and/or model biases. As such, the super-ensemble BHM posteriors were dominated by the model uncertainties.

During Prof. Nadia Pinardi's visit to CoRA in August, she, Berliner and Milliff worked with Dr. Jeremiah Brown (NWRA/CoRA) to identify two locations in the Rhodes Gyre region where time series of T and S profiles from SYS3a2, OPA and NEMO could be used as a new target process for the super-ensemble BHM. Time series of $T(z,t)$ and $S(z,t)$ have been provided to Prof. Berliner in preparation for new super-ensemble BHM experiments. The thinking is that inherent uncertainties in simulations and analysis of $T(z,t)$ and $S(z,t)$ will not be large enough to dominate posterior distributions in the new experiments to be run this month.

WORK COMPLETED AND RESULTS

Research program progress and plans were reviewed for ONR program managers in a presentation by Milliff at the ONR PI Meeting in Chicago in June 2009.

Publishing BHM-SVW development and forecast impacts

The careful production of the BHM-SVW equations and expressions for full conditional distributions, and the presentation of hyperprior specifications, are essential steps in preparing BHM-SVW for technology transfer to operational ocean forecast applications at MFS and elsewhere (e.g. NRL SSC Ocean Modeling Group). Together, these appendices comprise the necessary information to construct (and adapt as needed) the BHM-SVW used in our demonstration project with MFS.

Bonazzi et al. (2009) demonstrates impacts on MFS forecasts of uncertainty quantification based on realistic spread in surface wind forcing from BHM-SVW. Figure 1 provides a pictorial sense of the enhanced spread in ocean model response. The left hand panels depicts the SSH field at day 10 of ocean forecasts (i.e. day 10F) driven by members of the ECMWF ensemble forecast winds, and the right hand panel depicts the SSH fields on 10F as forced by 3 realizations of BHM-SVW. Superposed on each SSH field are 6 trajectories from idealized drifters released in each forecast at day 1F. The mesoscale response in the BHM-SVW case exhibits greater spread, as do the simulated trajectories. Ensemble spreads from traditional methodologies notoriously under-estimate uncertainties and often do not exhibit realistic variability. This failing is apparently obviated by BHM-SVW. Quantification of this improvement is documented in Bonazzi et al (2009).

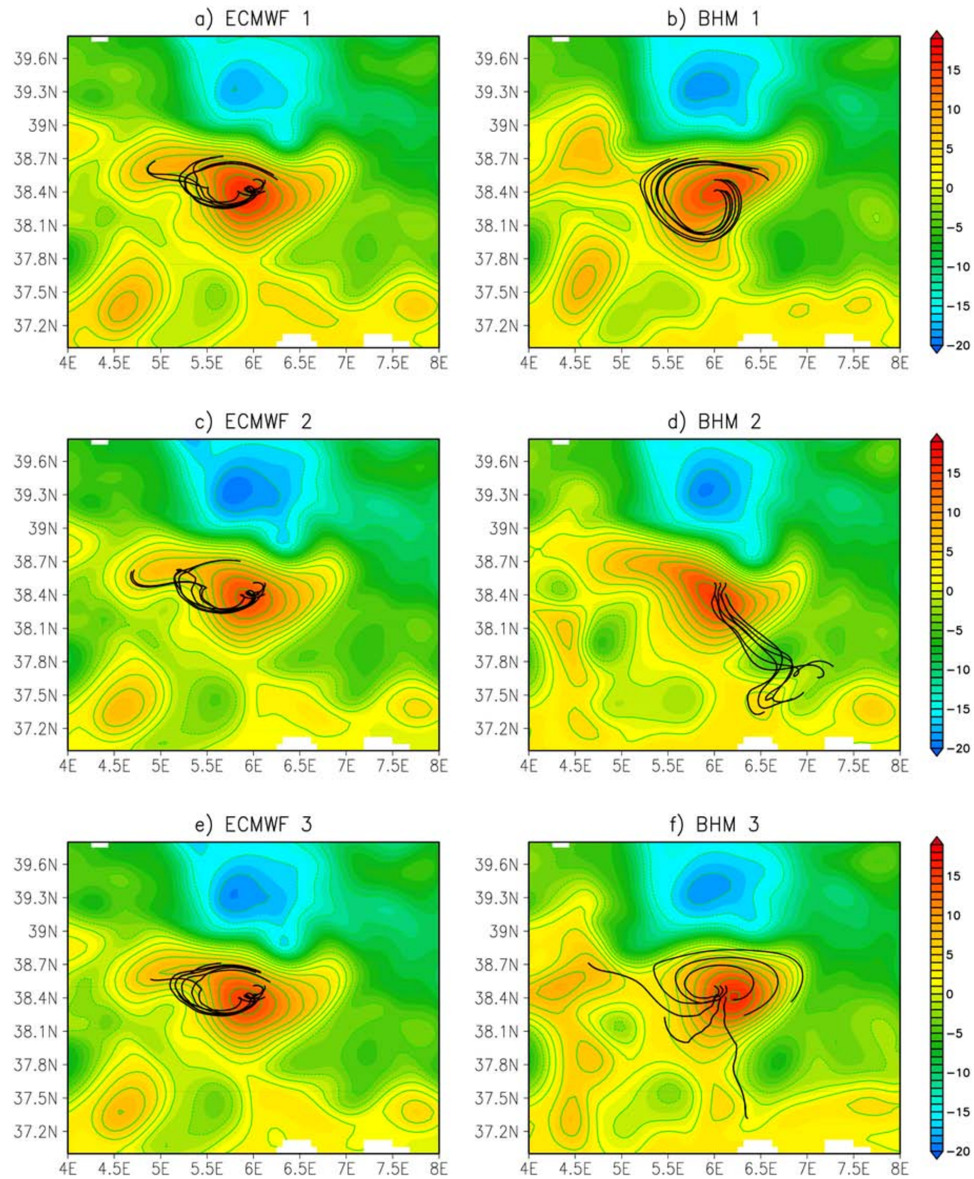

Figure 1. Sea-surface height anomaly at day 10 of ensemble ocean forecasts driven by (left) members of the ECMWF ensemble forecast winds and (right) members of the BHM-SVW posterior distribution. Color contours indicate SSH in cm, and black curves mark the trajectories of 6 idealized drifters released in the forecast fields at the forecast initial time. SSH variability on mesoscales, and spread in the trajectories are larger in the BHM-SVW ensemble. These results are quantified in Bonazzi et al. (2009). They are consistent with the localization and mesoscale focus in ocean forecast initial conditions documented in Milliff et al. (2009).

Impacts of the q-only time-dependent error covariance BHM

Figure 2 depicts a comparison of root-mean square differences in sea-level anomaly (SLA) with respect to an operational merged sea-level product based on altimeter data. Three forecast experiments are compared in MFS region 3, covering the same period, differing only in the form of the error covariance matrix used in the data assimilation step. The green line represents the operational system wherein the error covariance matrix is fixed in time. The black line is for a reforecast using time-dependent error covariance matrices from the BHM based on both ***d*** and ***q*** data stages, and the yellow line is for the recent ***q-only*** experiments. Surprisingly, time-dependent error covariance does not seem to improve the comparisons with SLA in either case. Further analyses (not shown) demonstrate that much of the uncertainty in SLA is coming from the error covariances for salinity (i.e. as opposed to temperature). New designs for the error covariance BHM are being designed with these findings in mind.

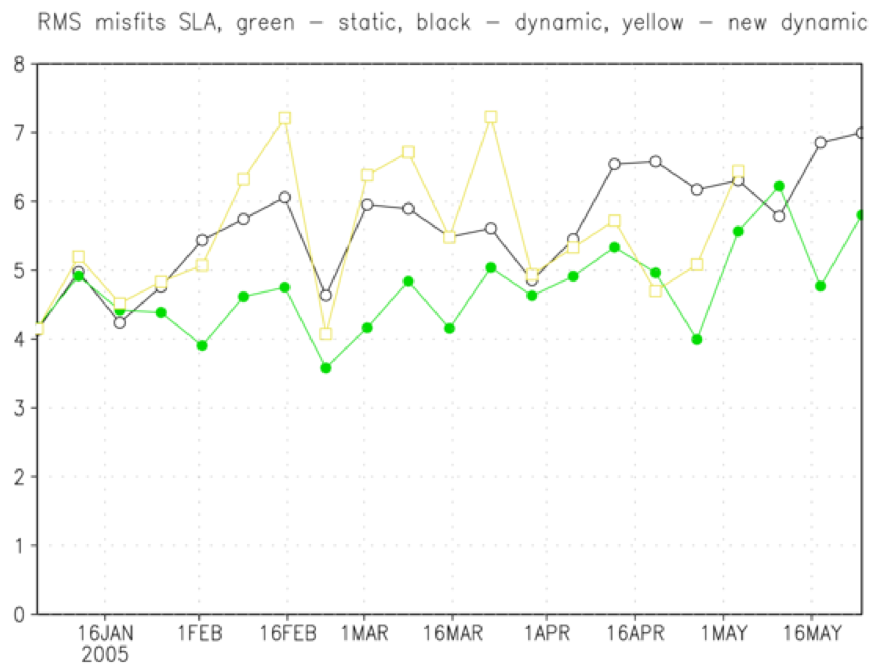

Figure 2. Sea-level anomaly (SLA) root-mean-square differences with respect to an operational satellite merged product for MFS model grid points in region 3 (Gulf of Lyon) for the period March-May 2005. Three reforecast experiments were run over the period, varying only the form of the error covariance matrix used in the data assimilation. The green line represents the operational system wherein the error covariance matrix is fixed in time. The black line represents a time-dependent error covariance that is the posterior mean output of the error covariance BHM based on misfit (d) data and anomaly (q) data. The yellow line corresponds to time-dependent error covariances from the BHM based on only the q data. The q-only reforecast experiment only ran to 1 May 2005.

Target field for the super-ensemble BHM

Figure 3 shows sample *T(z,t)* summaries for a point on the southwestern flank of the Rhodes Gyre in the Eastern Mediterranean Sea. The summary period shown runs from 1 February through 31 March 2007. *T(z,t)* and *S(z,t)* summaries for two points (the other location is on the northeastern flank of the Rhodes Gyre) have been created from the analysis (1 realization for each location) and from 11 simulations each of the OPA and NEMO development models at MFS. The OPA and NEMO simulations are forced by the ECMWF analysis and 10 realizations of the BHM-SVW. Similar summaries are provided for the same months in 2005 and 2006. The super-ensemble BHM will be used to hindcast the 2008 *T(z,t)* and *S(z,t)* summaries.

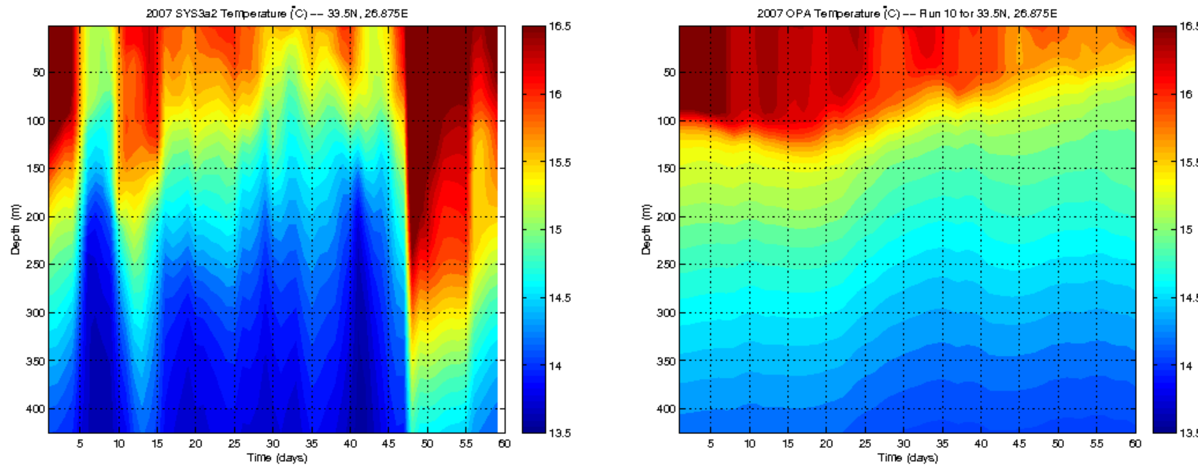

Figure 3. T(z,t) for a point at 26.875°E, 33.5°N (i.e. on the southwestern flank of the Rhodes Gyre) for the period 1 February through 31 March 2007. The left hand panel depicts T(z,t) from the MFS analysis (SYS3a2) and the right hand panel is from an OPA model simulation driven by one realization from BHM-SVW. An intermediate water formation event is evident in the analysis (left) around day 7. It is absent from the OPA simulation (right). Evidence of a mesoscale anticyclonic eddy appears in the analysis (left) near day 50. The mesoscale variability in the OPA simulation is much reduced vs. the analysis.

IMPACTS AND APPLICATIONS

Wikle and Milliff were invited speakers at a National Research Council workshop on "Uncertainty Management in Remote Sensing of Climate Data" in December 2008. The BHM-SVW was the topic of Milliff's talk.

Wikle, Milliff and Dr. Manuel Fiadeiro (ONR Pgm Mgr) are hosting a session on Probabilistic Modelling in Oceanography at the AGU 2010 Ocean Sciences Meeting in Portland, OR. Berliner and Pinardi are invited speakers.

RELATED PROJECTS

NSF-C433P-RFM-CORA-1433:
"Collaborative Research: Estimating Ecosystem Model Uncertainties in Pan-Regional Syntheses and Climate Change Impacts on Coastal Domains of the North Pacific Ocean". Co-PIs include Wikle and Milliff; Advisory Council includes Berliner. Period of Performance September 2008- August 2011.

ONR PO (in process):
"Bayesian Hierarchical Model Characterization of Model Error in Ocean Data Assimilation and Forecasts" Co-PIs include Berliner, Wikle and Milliff. Period of Performance to be determined (4yrs).

REFERENCES

Wikle, C.K., R.F. Milliff, D. Nychka, and L.M. Berliner, 2001: "Spatiotemporal hierarchical Bayesian modeling: Tropical ocean surface winds", *J. Amer. Stat. Assoc.,* **96***,* 382-397.

PUBLICATIONS

Berliner, L.M., and Y. Kim, 2008: "Bayesian design and analysis for super ensemble based climate forecasting", *J. Climate*, **21**, 1891-1910.

Bonazzi, A., N. Pinardi, S. Dobricic, R. Milliff, C. Wikle and L.M. Berliner, 2009: "Ocean Ensemble Forecasting, Part II: Mediterranean Forecast System Response", to be submitted October 2009.

Milliff, R.F., A. Bonazzi, C.K.Wikle, L.M. Berliner and N. Pinardi, 2009: "Ocean Ensemble Forecasting, Part I: Ensemble Mediterranean Winds from a Bayesian Hierarchical Model", to be submitted October 2009.